

Risk-Sensitive Mean-Field-Type Games with L^p -norm Drifts

Hamidou Tembine
 Learning & Game Theory Lab
 Division of Engineering
 New York University Abu Dhabi

September 22, 2014

Abstract

We study how risk-sensitive players act in situations where the outcome is influenced not only by the state-action profile but also by the distribution of it. In such interactive decision-making problems, the classical mean-field game framework does not apply. We depart from most of the mean-field games literature by presuming that a decision-maker may include its own-state distribution in its decision. This leads to the class of mean-field-type games. In mean-field-type situations, a single decision-maker may have a big impact on the mean-field terms for which new type of optimality equations are derived. We establish a finite dimensional stochastic maximum principle for mean-field-type games where the drift functions have a p -norm structure which weaken the classical Lipschitz and differentiability assumptions. Sufficient optimality equations are established via Dynamic Programming Principle but in infinite dimension. Using de Finetti-Hewitt-Savage theorem, we show that a propagation of chaos property with virtual particles holds for the non-linear McKean-Vlasov dynamics.

1 Introduction

Recently, there has been a renewed interest in optimization and game problems of mean-field type, where the performance functionals, drifts, diffusions, and jump coefficients depend not only on the state and the control but also on the probability distribution of state-control pair. Most formulations of mean-field type optimization in [5, 6, 8, 13, 16, 27] have been of risk-neutral type where the performance functionals are the expected values of stage-additive cost functions of Bolza or Mayer type. Not all behavior, however, can be captured by risk-neutral mean-field type optimizations. One way of capturing risk-averse and risk-seeking behaviors is by exponentiating the performance functional before expectation (see [14]). The objective of a risk-sensitive player is then to optimize

an exponentiated long-term loss. The risk-sensitive criterion is related to the robust control via relative entropic measures. As the risk-sensitive parameters vanish, one gets a risk-neutral maximum principle of mean-field-type.

1.1 On Mean-Field Games

There are several pioneer works on static and/or stationary mean-field games. Most of them are under different names such as global games, anonymous games, aggregative games, population games, large games, etc, but share lot of common features. Here we limit ourselves to the some pioneer works on dynamic mean-field games. One of the first works on mean-field games is [1]. Therein, the author proposes a game-theoretic model that explains why smaller firms grow faster and are more likely to fail than larger firms in large economies. The game is played over a discrete time space. The mean-field is the aggregate demand/supply which generates a price dynamics. The price moves forwardly, and the players react to the price and generate a demand and the firm a supply with associated cost, which regenerates the next price and so on. The author introduced a system of backward-forward system to find equilibria (see for example Section 4, equation D.1 and D.2 in [1]). The backward equilibrium equation is obtained as an optimality to the individual response, i.e., the value function associated with the best response to price, and the forward equation for the evolution of price. Therein, the consistency check is about the mean-field of equilibrium actions (population or mass of actions), that is, the equilibrium price solves a fixed-point system: the price regenerated after the reaction of the players through their individual best-responses should be consistent with the price they responded to. Following that analogy, a more general framework was developed in [2], where the mean-field equilibrium is introduced in the content of Markovian dynamic games with large number of decision-makers. A mean-field equilibrium is defined in page 4 of [2] by two conditions: (i) each generic player's action is best-response to the mean-field, and (ii) the mean-field is consistent and is exactly reproduced from the reactions of the players. This matching argument was widely used in the literature as it can be interpreted as a generic player reacting to an evolving mean-field object and at the same time the mean-field is formed from the contributions of all the players. The authors of [39] show that show how common noise can be introduced into the mean-field game model (so the mean-field distribution evolves stochastically) and extend the Jovanovic-Rosenthal existence theorem. The methodology developed in [1] and the subsequent series of papers [2, 39, 40, 41, 43, 44] share the following assumptions:

- (Big size) There is a large number of decision-makers, sometimes, infinite, or a continuum of decision-makers.
- (Anonymity) The index of the decision-maker does not affect the utility.
- (NonAtomicity) A single decision-maker has a negligible effect on the mean-field-term and on the utility.

Unfortunately, some of the above conditions appear to be very restrictive in terms of applications, and we explain below how to relax them via mean-field-type game theory.

1.2 Related works on Mean-Field-Type Game Theory

One decision-maker

A stochastic maximum principle (SMP) for the risk-sensitive optimal control problems for Markov diffusion processes with an exponential-of-integral performance functional was elegantly derived in [17] using the relationship between the SMP and the Dynamic Programming Principle (DPP) which expresses the first order adjoint process as the gradient of the value-function of the underlying control problem. This relationship holds only when the value-function is smooth (see Assumption (B4) in [17]). The approach of [17] was widely used and extended to jump processes in [19] and [20, 18], but still under this smoothness assumption. However, in many cases of interest, the value function is, in the best case, only continuous. Moreover, the relationship between the SMP and the DPP is unclear for non-Markovian dynamics and for mean-field type game problems where the Bellman optimality principle need to be extended. This calls for the need to find a risk-sensitive SMP and DPP for these cases. Djehiche et al. (2014, [22, 4]) have established a stochastic maximum principle for risk-sensitive mean-field-type control where the key mean-field term is the *mean state*. This means that the drift, diffusion, running cost and terminal cost functions depend on the state, the control and on the mean of state. Our work extends the results of [17] to risk-sensitive control problems for dynamics that are non-Markovian and of mean-field type. One important contribution of [22] is that the derivation of the SMP does not require any (explicit) relationship between the first-order adjoint process and a value-function of an underlying control problem. Using the SMP derived in [13], the approach is easily extended to the case where the mean-field coupling is in terms of the mean of the state and the control processes. In [45], we have extended the methodology to risk-sensitive mean-field-type control under partial observation which has interesting applications in risk-sensitive filtering problems including mean-field ensemble Kalman filtering, state tracking and other data assimilation algorithms in geosciences.

Two or more decision-makers

The first paper that deals with risk-sensitive games in a mean-field context is [21]. Therein, we have derived a verification theorem for a risk-sensitive mean-field game whose underlying dynamics is a Markov diffusion, using a matching argument between a system of Hamilton-Jacobi-Bellman (HJB) equations and the Fokker-Planck equation. This matching argument freezes the mean-field coupling in the dynamics, which yields a risk-sensitive HJB equation for the value-function. The mean-field coupling is then retrieved through the Fokker-Planck equation satisfied by the marginal law of the optimal state. The work in

[21] is fundamentally different than the present work. Therein, the mean-field term is frozen to be the equilibrium mean-field term and a single decision-maker cannot influence the mean-field-term. In the present work, we shall show that, when a single decision-maker has a non-negligible effect in the mean-field, the fundamental optimality equations are changed. In [24] we have analyzed risk-sensitive linear-exponentiated quadratic games of mean-field-type for which we have provided closed-form expressions using a novel risk-sensitive stochastic maximum principle derived in [22] which does not use the value function. It allows us, in particular, to work with the SMP equations in situations where the value function is not necessarily differentiable.

Substantial progress have been done in the last decade in mean-field games in the non-cooperative setup. However, very little is known about cooperative mean-field games. In [36] we have introduced cooperative mean-field type games in which the state dynamics and the payoffs depend not only on the state and actions but also on their probability measure. We establish a time-dependent payoff allocation procedure for coalitions of mean-field type. The allocated payoff considers not only fairness property but also the cost of making the coalition. Both time consistency and subgame perfectness solution concept equations are established.

1.3 Mean-field-type games: additional features

Risk-sensitive mean-field-type games [42] are fundamentally different than risk-sensitive mean-field games. In the mean-field game-theoretic models it is usually assumed that (i) very large number of players, (ii) players are indistinguishability (in the sense of the strategies, payoffs, state laws), (iii) individual contribution to the mean-field term is assumed to be negligible. In *mean-field-type games*, none of the assumptions (i)-(iii) is needed. Following [36], a mean-field-type game is defined as any game in which the payoff and/or state dynamics involve not only the state and action profiles and also the distribution of the state-action pair (or its marginals such as distribution of states and distribution of actions). Mean-field-type game theory is suitable for one, two or more players. A typical example is, a single decision-maker with mean-variance payoff. In mean-field-type games: (i) a single player can have a big influence on the mean-field term. A typical example is an Air Conditioning system which tries to reduce the variance of the temperature state with the respect the desired comfort temperature of the user. That the context, there is only one decision-maker, the user, who acts on the controller. The control variable is between $\{Heating, Cooling, Nothing\}$. Clearly, the control action has significant impact on the variance of the temperature. (ii) there is no need for players to be indistinguishable (see Section 3). (iii) there is no need to have large number (or infinite or continuum) of players. The mean-field-type game framework allows us to address more interesting real-world applications where the number of decision-makers may be large but still finite and include both von Neumann and non-von Neumann utility functions.

1.4 Novelty and Contribution

Our contribution can be summarized as follows. We start with one player risk-sensitive mean-field-type optimization where the state dynamics has L^p -norm structure, which is not differentiable. Our main motivation for considering this structure comes from its applications for the control of virus spread among interactive communities (networks) as observed in [25]. This allows us to consider other types of non-linear mean-field interactions that are not investigated in the literature of mean-field games. It also allow us to consider weakened Lipschitz conditions and non-differentiable drift coefficients. We show that the non-differentiability issue can be handled using weak derivatives or sub-differential set. We derive a stochastic maximum principle and a dual game variable which satisfies the risk-sensitive SMP whenever the associates weak derivatives make sense. In addition, a risk-sensitive DPP is provided in infinite dimension. We believe that the present paper is the first work that analyzes risk-sensitive mean-field-type games with the L^p -norm which is non-differentiable.

1.5 Structure of the paper

The paper is organized as follows. In Section 2, we present the model and state the main results for one player. Section 3 presents risk-sensitive mean-field type games with two or more players. We provide a dynamic programming principle in infinite dimension in subSection 3.2. Section 4 focuses on the control of virus spread among interactive communities (networks). Section 5 concludes the paper. For completeness, we provide in Appendix the de Finetti-Hewitt-Savage theorem and the existence and uniqueness proofs.

To streamline the presentation, we only consider the one-dimensional state case. The extension to the multidimensional case is by now straightforward. The norm is denoted with the index $\alpha \geq 1$ and p will be used for the adjoint process in the stochastic maximum principle. Also, it should be noted that our diffusion coefficient is control independent. More general state, control and mean-field dependent diffusions are carried out in [22]. Also the technique developed here can be easily extended to the jump-diffusion case using the works in [10, 12, 11].

2 Mean-field-type game with one risk-sensitive decision-maker

Let $T > 0$ be a fixed time horizon, $\alpha \geq 1$ and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a given filtered probability space on which a one-dimensional standard Brownian motion $B = \{B_s\}_{s \geq 0}$ is given, and the filtration $\mathbb{F} = \{\mathcal{F}_s, 0 \leq s \leq T\}$ is the natural filtration of B augmented by \mathbb{P} -null sets of \mathcal{F} . We consider the following risk-sensitive

problem :

$$\begin{cases} \bar{J}^\theta(u(\cdot)) \\ = \frac{1}{\theta} \log \left(E e^{\theta \left[\int_0^T f(t, x^u(t), m^u(t), u(t)) dt + h(x^u(T), m^u(T)) \right]} \right), \\ \inf_u \bar{J}^\theta(u) \\ \text{subject to} \\ dx^u(t) = \bar{b}(\cdot, t, x^u(t), m^u(t), u(t)) dt + \sigma(\cdot, t, x^u(t)) dB(t), \\ x^u(0) = x_0, m^u(t) := \mathcal{L}(x^u(t)), \end{cases} \quad (1)$$

where the state space is $\mathcal{X} = \mathbb{R}$, the term \bar{b} is distribution-dependent and has the special structure

$$\bar{b} = \left(\int_{y \in \mathcal{X}} |b|^\alpha(\cdot, t, x^u(t), y, u(t)) m^u(t, dy) \right)^{\frac{1}{\alpha}},$$

i.e., the L^α -norm of b with the respect to the measure $m^u(t, \cdot)$.

$$\bar{b}(t, x, m, u) : [0, T] \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times U \longrightarrow \mathbb{R},$$

$t \in [0, T]$, $x \in \mathbb{R}$, $m \in \mathcal{P}(\mathcal{X})$, $u \in U$. Notice that for $\alpha > 1$ the drift term \bar{b} is non-linear in the measure m .

$$b(\cdot, t, x, y, u) : [0, T] \times \mathcal{X}^2 \times U \longrightarrow \mathbb{R},$$

$$\sigma(\cdot, t, x) : [0, T] \times \mathcal{X} \longrightarrow \mathbb{R},$$

$m^u(t) := \mathcal{L}(x^u(t)) := P_{x^u(t)}$ is the probability law of the random variable $x^u(t)$. The parameter θ is the risk-sensitivity index of the player. The instantaneous cost function is

$$f(t, x, m, u) : [0, T] \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times U \longrightarrow \mathbb{R},$$

and the terminal cost function is

$$h(x, m) : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \longrightarrow \mathbb{R}.$$

The control strategy u is chosen by the decision-maker. An admissible control strategy u is an \mathbb{F} -adapted and L^{α^*} -integrable process with values in a non-empty subset U of \mathbb{R}^d . We denote the set of all admissible strategies of the player by \mathcal{U} .

Definition 1. *A mean-field-type game is a game in which the payoff and/or state dynamics involve not only the state-action profiles and also the distribution of the state-action pair (or its marginals such as distribution of states and distribution of actions).*

Example 1. *Problem (1) is a mean-field-type game with one decision-maker. The optimality equation of (1) is a nonstandard system from mean-field-type optimal control [28].*

Given an admissible strategy $u \in \mathcal{U}$ of the decision-maker (player), the state equation in (1) is a measure-dependent stochastic differential equation (SDE) with random coefficients.

In view of (1), up to a change of the parameter θ into $-\theta$, the optimization of \bar{J}^θ is the same as the following optimization problem

$$J^\theta(u(\cdot)) = E e^{\theta \left[\int_0^T f(t, x^u(t), m^u(t), u(t)) dt + h(x^u(T), m^u(T)) \right]}, \quad (2)$$

Any $\bar{u}(\cdot) \in \mathcal{U}$ satisfying

$$J^\theta(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}} J^\theta(u(\cdot)), \quad (3)$$

is called a risk-sensitive optimal strategy. The corresponding state process, solution of the SDE in (1), is denoted by $\bar{x}(\cdot) := x^{\bar{u}}(\cdot)$. The mean-field-type optimization problem that we are interested in, is to characterize the pair (\bar{x}, \bar{u}) solution of the problem (1). Let $\Psi_T = \int_0^T f(t, x(t), m^u(t), u(t)) dt + h(x(T), m^u(T))$. Then the risk sensitive loss functional is given by

$$\bar{J}^\theta = \frac{1}{\theta} \log J^\theta = \frac{1}{\theta} \log [E e^{\theta \Psi_T}].$$

When the risk-sensitive index θ is small, the loss functional \bar{J}^θ can be expanded as

$$E[\Psi_T] + \frac{\theta}{2} \text{var}(\Psi_T) + O(\theta^2),$$

where, $\text{var}(\Psi_T)$ denotes the variance of Ψ_T . If $\theta < 0$, the variance of Ψ_T , as a measure of risk, improves the performance, in which case the optimizer is called *risk seeker*. But, when $\theta > 0$, the variance of Ψ_T worsens the performance \bar{J}_θ , in which case the optimizer is called *risk averse*. The risk-neutral loss functional $E[\Psi_T]$ can be seen as a limit of risk-sensitive functional \bar{J}_θ when $\theta \rightarrow 0$. This is one of the reasons why this criterion attracted lots of attention. The criterion has also interesting connections with H_∞ -mean-field-type optimization. This is easily viewed from the Donsker-Varadhan formula:

$$\frac{1}{\theta} \log \left(\int e^{\theta \phi} d\nu \right) = \sup_{\mu \in \mathcal{P}(\Omega)} \left[\int \phi d\mu - \frac{1}{\theta} \tilde{H}(\mu | \nu) \right], \quad (4)$$

for any measurable bounded function ϕ on Ω , and ν a probability measure on Ω . Moreover, the supremum is *uniquely* achieved by the *imitative* Boltzmann-Gibbs distribution μ_* widely used in distributed strategic learning [26],

$$d\mu_* = \frac{e^{\theta \phi} d\nu}{\int_\Omega e^{\theta \phi} d\nu}.$$

The function $\tilde{H}(\cdot | \cdot)$ is the relative entropy from Ω to $\bar{\mathbb{R}}$ given by

$$\tilde{H}(\mu | \nu) := \int \log \left(\frac{d\mu}{d\nu} \right) d\mu = \int \frac{d\mu}{d\nu} \log \left(\frac{d\mu}{d\nu} \right) d\nu,$$

whenever $\mu \in \mathcal{P}(\Omega)$ is absolutely continuous with the respect to ν , otherwise we set $\tilde{H}(\mu|\nu) = +\infty$. The problem is

$$\inf_u \bar{J}^\theta(u) = \inf_u \sup_{\mu \in \mathcal{P}(\Omega)} \left[\mathbb{E}_\mu[\Psi_T] - \frac{1}{\theta} \tilde{H}(\mu|\mathbb{P}) \right], \quad (5)$$

2.1 Existence of solution to the state equation

We now focus on the well-posedness of the state dynamics.

Proposition 1. *If the functions b and σ are Lipschitz with the respect to (x, y) with Lipschitz constant $L > 0$ and*

$$\int_0^T \left(\int_y |b(t, 0, y, u)|^\alpha m(t, dy) \right)^{1/\alpha} dt < +\infty,$$

and then, the SDE in (1) admits a unique strong solution x^u in L^α . If in addition,

$$\int_0^T \left(\int_y |b(t, X, y, u)|^{2\alpha} m(t, dy) \right) dt < +\infty,$$

a.s. then $\forall \alpha \geq 2$,

$$\sup_n \sqrt{n} \left\{ \mathbb{E} \left(\sup_{t \leq T} |x_{i,n}(t) - \bar{x}_{i,n}(t)|^\alpha \right) \right\}^{\frac{1}{\alpha}} < +\infty,$$

where $x_{i,n}(t)$ a i -th particle state solution of

$$\begin{cases} dx_{i,n}^u(t) = \left(\frac{1}{n} \sum_{j=1}^n |b(\cdot, t, x_{i,n}^u(t), x_{j,n}^u(t), u(t))|^\alpha \right)^{\frac{1}{\alpha}} dt \\ + \sigma(\cdot, t, x_{i,n}^u(t)) dB_{i,n}(t), \\ \mathcal{L}(x_{i,n}^u(0)) = m_0, \end{cases} \quad (6)$$

and $\bar{x}_{i,n}(t)$ has the law of $x^u(t)$.

Proof. See Appendix. □

As we provide in Theorem 1 in Appendix, the mean-field convergence of the empirical measure $\frac{1}{n} \sum_{i=1}^n \delta_{x_{i,n}^u(t)}$ is by now a well-established result under de Finetti-Hewitt-Savage theorem. The issue here is to identify the limiting measure m with the particularity of the L^α -norm structure. We provide an example of mean-field-type SDE in cooperative dynamics.

Example 2 (Effect of mean-field in cooperative dynamics). *Consider the mean-field stochastic dynamics with drift*

$$\left\{ \int \left\| \sin(-x^3(t) + x(t)) - \mu \sin(x(t) - y) \right\|^\alpha m(t, dy) \right\}^{1/\alpha},$$

where $\mu > 0$, and constant diffusion coefficient $\sigma \in \mathbb{R}$. The first term in the drift ($\sin(-x^3(t) + x(t))$) is often replaced by a control action $u(t) \in [-1, 1]$ to get

$$dx(t) = \sigma dB(t) + \left\{ \int \|u(t) - \mu \sin(x(t) - y)\|^\alpha m(t, dy) \right\}^{1/\alpha} dt \quad (7)$$

This type of mean-field SDE models has been used to understand muscle contraction (see Section 5 in [23]). Other similar models have been widely studied in chemical kinetics, statistical mechanics and economics to capture cooperative behavior of a generic particle, oscillator or an agent.

Note that the presence of the measures $m^u(s)$, $0 \leq s \leq T$, in the loss function J^θ may cause time-inconsistency, in which case the Bellman's Principle with the state x is no longer valid and this motivates the use of the stochastic maximum principle (SMP) approach to get a finite dimensional framework. Note, however that, one can apply DPP where the state is infinite dimension $\mu(t, \cdot)$ as shown in Section 3.2 (see also [37]).

2.2 Stochastic Maximum Principle

We define the risk-neutral Hamiltonian associated with random variables $X \in L^\alpha(\Omega, \mathcal{F}, \mathbb{P})$ as follows. for $(p, q) \in \mathbb{R} \times \mathbb{R}$

$$H(t, X, m, u, p, q) := \bar{b}(t, X, m, u)p + \sigma(t, X)(t, X)q - f(t, X, m, u). \quad (8)$$

We also introduce the risk-sensitive Hamiltonian: for $\theta \in \mathbb{R}$ and $(p, q, \ell) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$,

$$H^\theta(t, X, m, u, p, q, \ell) := \bar{b}p + \sigma(q + \theta \ell p) - f. \quad (9)$$

We have $H = H^0$. The sign $(-f)$ is used here for the only purpose of having maximum principle instead of minimum principle and does not fundamentally change the methodology. Moreover, we denote

$$H_k^\theta(t) := p(t)\bar{b}_k(t) + (q + \theta \ell p)\sigma_k(t) - f_k(t), \quad (10)$$

for $k \in \{x, m\}$.

Note that even if b is differentiable, the drift coefficient \bar{b} which is

$$\left(\int_{y \in \mathcal{X}} b^\alpha(\cdot, t, x^u(t), y, u(t)) m^u(t, dy) \right)^{\frac{1}{\alpha}},$$

may not be differentiable at the points where $b(\cdot) = 0$. Denote by

$$\bar{b}_x(t) = \frac{\int_{y \in \mathcal{X}} b_x b^{\alpha-1}(\cdot, t, x^u(t), y, u(t)) m^u(t, dy)}{\bar{b}^{\alpha-1}}$$

if $\bar{b}^{\alpha-1}(x, m) > 0$. The case where $\bar{b}^{\alpha-1}(x, m) < 0$ is handled in a similar way. The differentiation with the respect to the measure m is considered in a Gateaux-derivative sense as in [28].

$$\lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} \bar{b}(\cdot, t, x, m + \epsilon d) = \int \bar{b}_m(\cdot, t, x, m)(\xi) d(d\xi).$$

Example 3. We provide Gateaux differentiation of $\|x\|_\alpha$ -based functions:

- *Mean state:* Let $f(\cdot, t, x, m) = \int y m(t, dy)$. Then,

$$\lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} f(\cdot, t, x, m + \epsilon d) = \int \xi d(d\xi).$$

The Gateaux-derivative with the respect to m is $f_m(\cdot, t, x, m)(\xi) = \xi$. Then $f_m(\cdot, t, \xi, m)(x) = x$, $\partial_x[f_m(\cdot, t, \xi, m)(x)] = 1$. It is therefore clear that $\partial_x[f_m(\cdot, t, \xi, m)(x)] = 1 \neq 0 = \partial_m[f_x]$.

- *Square of the mean:*

Let $g(\cdot, t, x, m) = \frac{1}{2}(\int y m(t, dy))^2$. Then,

$$\lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} g(\cdot, t, x, m + \epsilon d) = \bar{m} \int \xi d(d\xi).$$

$g_m(\cdot, t, x, m)(\xi) = \xi \bar{m}$. Hence, $g_m(\cdot, t, \xi, m)(x) = x \bar{m}$, and $\partial_x g_m(\cdot, t, \tilde{X}, m)(x) = \bar{m}$.

- *Second moment:* If $g(\cdot, t, x, m) = \frac{1}{2} \int y^2 m(t, dy)$ then,

$$\lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} g(\cdot, t, x, m + \epsilon d) = \frac{1}{2} \int \xi^2 d(d\xi).$$

$g_m(\cdot, t, x, m)(\xi) = \frac{1}{2} \xi^2$. Hence, $g_m(\cdot, t, \xi, m)(x) = \frac{1}{2} x^2$. $\partial_x g_m(\cdot, t, \tilde{X}, m)(x) = x$.

- *α -th moment:* $g(\cdot, t, x, m) = \int y^\alpha m(t, dy)$ then, $\lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} g(\cdot, t, x, m + \epsilon d) = \int \xi^\alpha d(d\xi)$. Thus, $g_m(\cdot, t, \xi, m)(x) = x^\alpha$, and $\partial_x g_m(\cdot, t, \tilde{X}, m)(x) = \alpha x^{\alpha-1}$.

- *α -norm:* $g(\cdot, t, x, m) = (\int |y|^\alpha m(t, dy))^{1/\alpha} = m_\alpha^{1/\alpha}$ then,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0_+} \frac{d}{d\epsilon} g(\cdot, t, x, m + \epsilon d) \\ = \frac{1}{\alpha} \left[\int |\xi|^\alpha m(d\xi) \right]^{\frac{1}{\alpha}-1} \left[\int |\xi|^\alpha d(d\xi) \right]. \end{aligned}$$

Thus, $g_m(\cdot, t, \xi, m)(x) = \frac{x^\alpha}{\alpha m_\alpha^{\alpha-1}}$, and $\partial_x g_m(\cdot, t, \tilde{X}, m)(x) = \frac{x^{\alpha-1}}{m_\alpha^{\alpha-1}}$.

- L^α -normed drift: We compute the Gateaux-derivative of the L^α -normed drifts: $\bar{b}_m(., t, x, m)(\xi) := \frac{b^\alpha(., t, x, \xi)}{\alpha \bar{b}^{\alpha-1}}$. By changing variables, one has $\bar{b}_m(., t, \xi, m)(x) := \frac{b^\alpha(., t, \xi, x)}{\alpha \bar{b}^{\alpha-1}(t, \xi, m)}$. We differentiate with the respect to x to get:

$$\partial_x \bar{b}_m(., t, \xi, m)(x) = \frac{b^{\alpha-1}(., t, \xi, x) b_y(., t, \xi, x)}{\bar{b}^{\alpha-1}(t, \xi, m)}. \quad (11)$$

$E[L \partial_x \bar{b}_m(., t, X, m)(x)] = E[L \frac{b^{\alpha-1}(., t, X, x) b_y(., t, X, x)}{\bar{b}^{\alpha-1}(t, X, m)}] := \tilde{E}[\tilde{L} \frac{b^{\alpha-1}(., t, \tilde{X}, x) b_y(., t, \tilde{X}, x)}{\bar{b}^{\alpha-1}(t, \tilde{X}, m)}]$, where the notation \tilde{E} denotes the expectation with the respect to the variables with \tilde{X} which is an copy of X . We now replace the argument x by X to get $\tilde{E}[\tilde{L} \partial_x \bar{b}_m(., t, \tilde{X}, m)(X)] = \tilde{E}[\tilde{L} \frac{b^{\alpha-1}(., t, \tilde{X}, X) b_y(., t, \tilde{X}, X)}{\bar{b}^{\alpha-1}(t, \tilde{X}, m)}]$. If $\alpha = 1$, one gets $\tilde{E}[\tilde{L} b_y(., t, \tilde{X}, X)]$.

We now introduce the first order adjoint processes involved in the risk-sensitive SMP. The (risk-sensitive) first order adjoint equation is the following backward SDE of mean-field type:

$$\begin{cases} dp(t) = - \left\{ H_x^\theta(t) + \frac{1}{v^\theta(t)} E[v^\theta(t) \partial_x H_m^\theta(t)] \right\} dt \\ \quad + q(t)(-\theta \ell(t) dt + dB_t), \\ dv^\theta(t) = \theta \ell(t) v^\theta(t) dB_t, \\ v^\theta(T) = \phi^\theta(T), \\ p(T) = -h_x(T) - \frac{1}{\phi^\theta(T)} E[\phi^\theta(T) \partial_x h_m(T)]. \end{cases} \quad (12)$$

where,

$$\phi^\theta(T) := e^{\theta[h(\bar{x}(T), \bar{m}(T)) + \int_0^T f(t, \bar{x}(t), \bar{m}(t), \bar{u}(t)) dt]}. \quad (13)$$

Note that the Hamiltonian terms in (12) are evaluated at the optimal state and optimal control $(\bar{x}(\cdot), \bar{u}(\cdot))$, i.e., $H_k^\theta(t, \bar{x}(t), \bar{m}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \ell(t))$, $k \in \{x, m\}$.

Lemma 1 ([9, 8]). *Consider the following mean-field backward SDE*

$$\begin{aligned} p(t) = p(T) + \int_t^T \tilde{E}[\hat{f}(s, \tilde{p}(s), \tilde{q}(s), p(s), q(s))] ds \\ - \int_t^T q(s) dB(s), \end{aligned}$$

where $p(T)$ is a progressively measurable, square integrable random variable. Let $\hat{f}(t, ., ., ., .)$ be Lipschitz for all time $t \in [0, T]$ and $t \mapsto f(t, 0, 0, 0, 0)$ be square integrable over $[0, T]$. Then, the mean-field backward SDE has a unique adapted solution satisfying

$$E \left[\sup_{t \in [0, T]} |p(t)|^2 + \int_0^T |q(t)|^2 dt \right] < \infty. \quad (14)$$

Note that, by choosing $\hat{f}(t, \tilde{p}(t), \tilde{q}(t), p(t), q(t)) = a_0(t, \cdot) + a_1(t, \cdot)\tilde{p}(t) + a_2(t, \cdot)\tilde{q}(t) + a_3(t, \cdot)p + a_4(t, \cdot)q(t)$ where $a_i(t, \cdot)$ are measurable bounded coefficient functions, one gets a backward equation in the form of the adjoint equations.

Proposition 2. *If the functions b, σ, f, h , are twice continuously differentiable with respect to (x, m) and b, σ, f, h and all their first order derivatives with respect to (x, m) are continuous in (x, m, u) , and bounded then (12) admits an \mathbb{F} -adapted solution $(\bar{p}, \bar{q}, v^\theta, \ell)$ such that*

$$E \left[\sup_{t \in [0, T]} |\bar{p}(t)|^2 + \sup_{t \in [0, T]} |v^\theta(t)|^2 + \int_0^T (|\bar{q}(t)|^2 + |\ell(t)|^2) dt \right] < \infty. \quad (15)$$

In addition, if $b > 0$ then (12) admits a unique \mathbb{F} -adapted solution.

Proof. Under the assumptions of Proposition 2, the processes p, q solve a backward SDE coupled with the process v^θ . Moreover, these equations can be transformed into linear SDEs of mean-field-type, which involves $(p, q, E[p], E[q])$ and random coefficients involving $(x, L^\theta := \frac{v^\theta}{E[\phi^\theta(T)]}, l)$. We now check that the coefficients of the linear SDEs does not blow-up within the horizon $[0, T]$. For the functions σ, f, h , and their derivatives the boundedness follow from the assumption. However, it is not immediate for the drift coefficient \bar{b} . We recall that $\bar{b}_x, \bar{b}_m, \bar{b}_{xm}$ are not clearly defined at the point where $b(\cdot, t, x, \cdot)$ is zero. Since $\alpha \geq 1$, we replace these terms by any representation in the sub-differential set. All terms are bounded by $M = \sup_{t \in [0, T]} \sup |b_x(t, \cdot)|$. The process $v^\theta(t)$ is almost surely bounded. Then, for each direction chosen in the sub-differential, the assumptions of Lemma 1 are fulfilled and hence, the existence of solution to the first order risk-sensitive adjoint equations follows. Moreover, if $b > 0$ then the denominator does not vanish and \bar{b}_x and $\partial_x \bar{b}_m$ are (uniquely) well-defined, and bounded by M . Using Lemma 1 again we get existence and uniqueness of solution. \square

Note that the boundedness and differentiation conditions can be weakened by using the techniques developed in [7, 15]. The following Proposition is the stochastic maximum principle for Problem (1).

Proposition 3. (Risk-sensitive maximum principle) *Let the Assumptions of Proposition 2 hold. If $(\bar{x}(\cdot), \bar{u}(\cdot))$ is an optimal solution of the risk-sensitive control problem (1)-(3), then there are two pairs of \mathbb{F} -adapted processes (v^θ, ℓ) , (\bar{p}, \bar{q}) that satisfy (12)-(15) respectively, such that*

$$\begin{aligned} & H^\theta(t, \bar{x}(t), \bar{m}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \ell(t)) \\ &= \max_u H^\theta(t, \bar{x}(t), \bar{m}(t), u, \bar{p}(t), \bar{q}(t), \ell(t)), \end{aligned}$$

for almost every $t \in [0, T]$ and \mathbb{P} -almost surely.

Proof. To prove the SMP, we use a logarithmic transformation and follows similar steps as in [22]. \square

Below we provide an explicit representation of the process of the SMP via a dual approach and partial differential equations of mean-field type.

3 Mean-Field-Type Games: two or more risk-sensitive players

We now consider two or more risk-sensitive players. The risk-sensitivity index of player i is θ_i . The best response to u_{-i}, m is the following problem :

$$\begin{cases} \bar{J}_i^\theta(u_i(\cdot), u_{-i}(\cdot)) = \\ \frac{1}{\theta_i} \log \left(E e^{\theta_i \left[\int_0^T f_i(t, x^u(t), m^u(t), u(t)) dt + h_i(x^u(T), m^u(T)) \right]} \right), \\ \inf_{u_i} \bar{J}_i^\theta(u_i, u_{-i}) \\ \text{subject to} \\ dx^u(t) = \bar{b}(t, x^u(t), m^u(t), u(t)) dt + \sigma(t, x^u(t)) dB(t), \\ x^u(0) = x_0, \\ m^u(t) := \mathcal{L}(x^u(t)), \end{cases} \quad (16)$$

where u_{-i} denotes $(u_1, \dots, u_i, u_{i+1}, \dots, u_n)$, $n \geq 2$. and by abuse of notation, $u = (u_i, u_{-i})$. Note that we cannot impose indistinguishability of the players since θ_i and the objectives f_i, h_i may be different across the players.

3.1 Main Result

We now present the key results of the paper. The risk-sensitive game with cost J_i^θ solves a system of risk-sensitive HJB equations

$$\begin{aligned} 0 = & V_{i,t}(\mu) + \int_u \inf_u \left[\bar{b} \partial_x V_{i,\mu}(\mu)(t, x, z) + f \partial_z V_{i,\mu}(\mu)(t, x, z) \right. \\ & \left. + \frac{\sigma^2}{2} \partial_{xx} V_{i,\mu}(\mu)(t, x, z) \right] \mu(t, dx, dz), \end{aligned} \quad (17)$$

which is a partial differential equation with state μ (in infinite dimension). If we denote $v_i^*(t, x, z) := V_{i,\mu}(t, \mu)(t, x, z)$ as a dual function (because of the Gateaux derivative with the respect to μ) then (v_1^*, \dots, v_n^*) is in finite dimension and solves the dual system

$$\begin{aligned} 0 = & v_{i,t}^*(t, x, z) + v_{i,z}^*(t, x, z) H_i^*(t, x, \frac{v_{i,x}^*(t, x, z)}{v_{i,z}^*(t, x, z)}, m) \\ & + \frac{1}{2} \sigma^2 v_{i,xx}^*(t, x, z) + \\ & \int_{\tilde{w}} v_{i,z}^*(t, \tilde{w}) H_{i,m}^*(t, \tilde{x}, \frac{v_{i,x}^*(t, \tilde{w})}{v_{i,z}^*(t, \tilde{w})}, m)(x) \mu(t, d\tilde{w}) \\ & i \in \{1, 2, \dots, n\} \end{aligned} \quad (18)$$

where H_i is the Hamiltonian, $u_i \in \arg \min H_i$, $\tilde{w} = (\tilde{x}, \tilde{z})$, and $m(t, x) = \int \mu(t, x, dz_1 \dots dz_n)$, and $\mu(t, \cdot)$ solves the Kolmogorov equation in which u is replaced by the optimal strategies (u_1, u_2, \dots, u_n) . Then $(p_i^*, q_i^*, \eta_i^*, l_i^*)$ solves the (risk-sensitive) stochastic maximum principle system given by:

$$\begin{cases} dp_i^* = - \left[H_{i,x}^* + \frac{1}{\eta_i^*} E[\eta_i^* \partial_x H_{i,m}^*] + \sigma_x \left(\frac{\sigma v_{i,xx}^*}{\eta_i^*} \right) \right. \\ \quad \left. - p_i^* (l_i^*)^2 + \frac{\sigma l_i^* v_{i,xx}^*}{\eta_i^*} \right] + (-p_i^* l_i^* + \frac{\sigma v_{i,xx}^*}{\eta_i^*}) dB_i \\ d\eta_i^* = \eta_i^* l_i^* dB_i, \\ \eta_i^*(T) = \theta_i e^{\theta_i [z_i(T) + h_i(x(T), m(T))]}, \\ q_i = (-p_i^* l_i^* + \frac{\sigma v_{i,xx}^*}{\eta_i^*}) \\ p_i^*(T) = h_{i,x}(T) + \frac{1}{\phi_i^\theta(T)} E[\phi_i^\theta(T) \partial_x h_{i,m}(T)], \\ i \in \{1, 2, \dots, n\} \end{cases} \quad (19)$$

where,

$$\phi_i^\theta(T) := e^{\theta_i [h_i(\bar{x}(T), \bar{m}(T)) + \int_0^T f_i(t, \bar{x}(t), \bar{m}(t), \bar{u}(t)) dt]}. \quad (20)$$

3.2 Dynamic programming for risk-sensitive mean-field-type games

We establish a dynamic programming principle in infinite dimension. We first write the objectives as a function of the infinite dimensional state μ which satisfies the Fokker-Planck-Kolmogorov forward equation

$$\mu_t = -\partial_x [\bar{b}\mu] - \partial_z (f\mu) + \frac{1}{2} \partial_{xx} (\sigma^2 \mu) =: \tilde{b}, \quad (21)$$

with the initial distribution $\mu(0, dx, dz) = m_0(dx) \delta_0(dz)$. The advantage now is that $\mu(\cdot)$ is a deterministic object the cost can be rewritten in a deterministic manner as

$$J_i^\theta = \int \mu(T, dx, dz) e^{\theta_i (z_i + h_i(x, \int_{\bar{z}} \mu(T, \cdot, d\bar{z})))}.$$

This a terminal cost in the sense it is evaluated only at $\mu(T, \cdot)$. Since there is no running cost, one can write directly the HJB equation using classical calculus of variations for

$$V_i(t, \mu(t, \cdot)) = \inf_u \int \mu(T, dx, dz) e^{\theta_i (z_i + h_i(x, \int_{\bar{z}} \mu(T, \cdot, d\bar{z})))},$$

starting from $\mu(t, \cdot)$ at time t :

$$0 = V_{i,t} + \inf_u [\langle \tilde{b}, V_{i,\mu} \rangle],$$

where

$$\langle \tilde{b}, V_{i,\mu} \rangle = \int V_{i,\mu}(\mu)(\tilde{x}, \tilde{z}) \tilde{b}(\tilde{x}, \tilde{z}) d\tilde{x} d\tilde{z} \quad (22)$$

$$= - \int V_{i,\mu}(\mu)(x, z) \{ \partial_x [\bar{b}\mu] + \partial_z (f_i \mu) \} \quad (23)$$

$$+ \int V_{i,\mu}(\mu)(x, z) \frac{1}{2} \partial_{xx} (\sigma^2 \mu) dx dz \quad (24)$$

$$= \int \left[\bar{b} \partial_x V_{i,\mu} + f_i \partial_z V_{i,\mu} + \frac{\sigma^2}{2} \partial_{xx} V_{i,\mu} \right] \mu(t, dx, dz)$$

As we can see, the required working state for player i is (x, z_i) , therefore the partial derivatives of V_i with respect to z are only considered for z_i . The risk-sensitive HJB minimum principle yields

$$0 = V_{i,t} + \int \inf_u \left[\bar{b} \partial_x V_{i,\mu} + f_i \partial_z V_{i,\mu} + \frac{\sigma^2}{2} \partial_{xx} V_{i,\mu} \right] \mu(t, dx, dz), \quad (25)$$

This is an infinite dimensional PDE on (t, μ) . Below we provide a simpler optimality equation (i.e., the state will be in finite dimension) by setting $p_i^* = \frac{\partial_x V_{i,\mu}}{\partial_z V_{i,\mu}}$ and $V_{i,\mu}(\mu)(t, x, z) = v_i^*(t, x, z)$. Differentiating (25) with the respect to μ one gets

$$0 = \partial_t V_{i,\mu}(\mu)(t, x, z) + \hat{H}_i(\mu) + \int \hat{H}_{i,\mu}(\mu)(t, \tilde{x}, \tilde{z}) \mu(t, d\tilde{x}, d\tilde{z}) \quad (26)$$

where

$$\hat{H}_i = \bar{b} \partial_x V_{i,\mu} + f_i \partial_z V_{i,\mu} + \frac{\sigma^2}{2} \partial_{xx} V_{i,\mu} \quad (27)$$

where $H_i^*(t, x, p^*, m) = \inf_u [\bar{b} p_i^* + f_i]$.

Definition 2. The function $v_i^*(t, x, z) := V_{i,\mu}(\mu)(t, x, z)$ is called *Dual Function* associated with the best response value of player i .

$v_i^*(t, x, z) := V_{i,\mu}(\mu)(t, x, z)$ solves the PDE (18) where the state is now reduced to (x, z) which is in finite dimension.

Below we show that if there exists a dual function (in the sense of weak derivatives) then its weak derivatives provide a risk-sensitive SMP.

3.3 Dual functions associated with the best response values

In a risk-neutral setting, Bensoussan et al. have established in [28] a partial differential equation as a necessary condition for optimality under smoothness

assumption. We apply the methodology to the risk-sensitive case. The basic idea consists to write the optimality inequality as $J_i^\theta(u_i + \epsilon d_i, u_{-i}, m^{u_i + \epsilon d_i, u_{-i}}) - J_i^\theta(u, m^u) \geq 0$ for the cost functional J^θ . By introducing the auxiliary state z such that $dz_i = f_i(\cdot) dt$, $z_i(0) = 0$, the risk-sensitive game problem is transformed into mean-field-type game problem without running cost. The terminal cost is $e^{\theta_i[z_i(T) + h_i(x(T), m(T))]}$. Since the state is augmented to be (x, z) , the first order adjoint process becomes (p_{1i}, p_{2i}, q_i) and the unmaximized Hamiltonian is $\bar{b}p_{1i} + fp_{2i} + \sigma q_i$. Since the diffusion does not depend on the control u , the derivative of this term with the respect to u can be written as $p_{2i}\partial_u[\bar{b}\frac{p_{1i}}{p_{2i}} + f_i]$, whenever $p_{2i} \neq 0$.

Let $v_i^*(t, x, z)$ be the dual function defined above, satisfying (18) where $\mu = \mathcal{L}(x^u(t), z^u(t))$, the x -marginal is $m(t, \cdot) = \int \mu(t, \cdot, dz)$ and

$$H_i^*(t, x, p^*, m) = \inf_u [\bar{b}p_i^* + f_i].$$

The terminal condition is

$$\begin{aligned} v_i^*(T, x, z) &= e^{\theta_i(z_i + h_i(x, m(T)))} \\ &+ \theta_i \int h_{i,m}(\tilde{x}, m(T))(x) e^{\theta_i(\tilde{z}_i + h_i(\tilde{x}, m(T)))} \mu(T, d\tilde{x}, d\tilde{z}). \end{aligned}$$

If $(v_i^*)_i$ solves the dual equation (18) then

$$\begin{aligned} \lim_{\epsilon \rightarrow 0+} \frac{d}{d\epsilon} J_i^\theta &= \\ \int_{(t,x) \in [0,T] \times \mathcal{X}} v_{i,z}^* H_{i,u}(t, x, m, u, \frac{v_{i,x}^*}{v_{i,z}^*}, \frac{\sigma v_{i,xx}^*}{2v_{i,z}^*}) \mu(t, x, z) dx dt \end{aligned}$$

and $H_{i,u} = 0$ for interior optimal control u . Let v_i be a function in the Lebesgue space $L^1(I)$, with $I = [a, b]$ a compact interval of \mathbb{R} , $a < b$. We say that $w \in L^1(I)$ is a "weak derivative" of v if,

$$\int_I v(t) \varphi'(t) dt = - \int_I w(t) \varphi(t) dt,$$

for all infinitely differentiable functions φ with $\varphi(a) = \varphi(b) = 0$.

Equation (18) is an interesting partial differential equation. Indeed, if there is a solution $v^*(t, x, z)$ to (18) that is three times weakly differentiable then the partial weak derivatives of $v^*(t, x, z)$ solves the risk-sensitive SMP (12). Below we identify explicitly the processes solution to the risk-sensitive SMP. Let $p_i^*(t) = \frac{v_{i,x}^*(t, x^u(t), z^u(t))}{v_{i,z}^*(t, x^u(t), z^u(t))} = \frac{\partial_x V_{i,\mu}(t, x^u(t), z^u(t))}{\partial_z V_{i,\mu}(t, x^u(t), z^u(t))}$ where the derivatives are taken in a distribution sense (weak derivative). Then the process p^* evaluated at the optimal trajectory solves the backward SDE:

$$\begin{aligned} dp_i^* &= - \left[H_{i,x}^* + \frac{1}{\eta_i^*} E[\eta_i^* \partial_x H_{i,m}^*] \right] dt \\ &\quad - [\sigma_x(q_i^* + p_i^* l_i^*) + q_i^* l_i^*] dt + q_i^* dB, \end{aligned} \quad (28)$$

where

$$q_i^* = -p_i^* l_i^* + \frac{\sigma v_{i,xx}^*}{\eta_i^*}, \quad l_i^* = \sigma \partial_x [\log \eta_i^*],$$

$$\eta_i^* := v_{i,z}^*(t, x, z) = \partial_z V_{i,\mu}(\mu)(t, x, z),$$

We impose a strong smoothness on v^* and show that η^* solves a backward SDE similar to the one satisfied by v^θ .

Proposition 4. $\eta_i^* := v_{i,z}^*(t, x, z)$ solves the backward SDE:

$$d\eta_i^* = \eta_i^* l_i^* dB, \quad \eta_i^*(T) = \theta_i e^{\theta_i [z_i(T) + h_i(x(T), m(T))]}.$$
 (29)

Proof. By Ito's formula, we have

$$\begin{aligned} d\eta_i^* &:= dv_{i,z}^*(t, x, z) \\ &= [v_{i,zt}^* + v_{i,xz}^* \bar{b} + v_{i,zz}^* f_i + \frac{1}{2} \sigma^2 v_{i,xxz}^*] dt + \sigma v_{i,xz}^* dB. \end{aligned}$$
 (30)

From (18) it is clear that the partial (weak) derivative of the integral term with the respect to z is zero (because it does not depend on z). We differentiate (18) to get

$$v_{i,zt}^* + \partial_z \left[v_{i,z}^* H_i^*(t, x, \frac{v_{i,x}^*}{v_{i,z}^*}, m) \right] + \frac{1}{2} \sigma^2 v_{i,xx}^* = 0,$$

which means that the drift term is $v_{i,tz}^* + v_{i,xz}^* \bar{b} + v_{i,zz}^* f + \frac{1}{2} \sigma^2 v_{i,xxz}^* = 0$. Thus, $d\eta_i^* = \sigma v_{i,xz}^* dB$. Lets compute the diffusion coefficient more explicitly:

$$\sigma v_{i,xz}^* = \frac{v_{i,z}^*}{v_{i,z}^*} \sigma v_{i,xz}^* = \eta_i^* l_i^*.$$
 (31)

Hence, one gets

$$d\eta_i^* = \eta_i^* l_i^* dB, \quad \eta_i^*(T) = \theta_i e^{\theta_i [z_i(T) + h_i(x(T), m(T))]}.$$

□

Proposition 5. The function $(\eta_1^*, \dots, \eta_n^*)$ solves the partial differential equation:

$$0 = \eta_{i,t}^* + \eta_{i,x}^* \bar{b} + \eta_{i,z}^* f_i + \frac{1}{2} \sigma^2 \eta_{i,xx}^* \quad (32)$$

whenever these derivatives make sense. Moreover η_i^* has a constant sign and has the same sign as θ_i .

Proof. This follows from a weak derivative with the respect to z_i in Eq. (18). □

Note that the function m in Proposition 5 is the marginal of μ with the respect to x and μ solves the Fokker-Planck-Kolmogorov forward equation with drifts (\bar{b}, f) and diffusion coefficient $(\sigma, 0)$:

$$\mu_t + \partial_x [\bar{b} \mu] + \partial_z (f \mu) - \frac{1}{2} \partial_{xx} (\sigma^2 \mu) = 0, \quad (33)$$

Table 1: Internet attacks over the globe. An increase of around 13.02%

Online Attacks								
North America			Europe		Asia		Australia	
	New York	617 150	Germany	571255	India	2 353 001	Australia	705594
	Virginia	452916	Romania	226018	China	772447		
	Illinois	401766	Sweden	138543	Bangladesh	106102		
	California	343627						
	Texas	309426						
	Ohio	110137						
	Florida	106779						

$\mu(0, dx, dz) = m_0(dx)\delta_0(dz)$. In view of Proposition 4, $(\bar{p}_i, \bar{q}_i, v_i^\theta, \ell_i) = (-p_i^*, q_i^*, \frac{\eta_i^*}{\theta_i}, \theta_i l_i^*)$ solves the risk-sensitive SMP (12) and

$$\begin{aligned} dp_i^* = & - \left[H_{i,x}^* + \frac{1}{\eta_i^*} E[\eta_i^* \partial_x H_{i,m}^*] + \sigma_x \left(\frac{\sigma v_{i,xx}^*}{\eta_i^*} \right) \right] dt \\ & - \left[-p_i^* (l_i^*)^2 + \frac{\sigma l_i^* v_{i,xx}^*}{\eta_i^*} \right] dt + (-p_i^* l_i^* + \frac{\sigma v_{i,xx}^*}{\eta_i^*}) dB, \end{aligned} \quad (34)$$

$$\begin{aligned} p_i^*(T) = & h_{i,x}(x(T), m(T)) + \\ & \frac{\tilde{E} \left[e^{\theta_i(z_i(T) + h(\tilde{x}(T), m(T)))} \partial_x h_{i,m}(\tilde{x}(T), m(T))(x(T)) \right]}{e^{\theta_i(z_i(T) + h_i(x(T), m(T)))}} \end{aligned}$$

The function $v^*(t, \cdot) = \partial_\mu V$ is not the value function in the sense of Bellman because of the presence of the term $E[H_{i,m}]$ in Eq. (18). $v^*(t, \cdot)$ is the adjoint function (dual function) associated to mean-field-type best-response problem. Interestingly, in the mean-field free case, i.e., when $h_{i,m} = 0, f_{i,m} = 0, b_{i,m} = 0$, the dual function $(v_1^*(t, \cdot), \dots, v_n^*(t, \cdot))$ coincides with the best-response value function of the risk-sensitive game problem with augmented state (x, z) .

4 Virus Spread over an evolving network

WiFi network security has gained significant attention in research and industrial communities as a result of the global connectivity provided by the Internet. This has led to a variety of traditional defense mechanisms ranging from cryptography, firewalls, antivirus software, to intrusion detection systems. Table 1 displays a sample number of network attacks by major geographic region (State or Country) with more than 100 000 attacks. See [38] for more details on real time web attack monitoring.

A virus that spreads through WiFi networks as effectively as a human cold moves through cities, airports, public transport areas has been explored recently. The virus can travel between WiFi networks via Access Points (APs)

that connect households and businesses to WiFi networks. It can also propagate through femto cell and small cell networks. We denote by x the state of the entire network. x could represent the number of access points that can be reached with infected relays (hotspots) at a specific period of the day. Since the number of access points that are active is highly stochastic and the number of nodes in the network is time-varying, x is a random process. We do not consider a mass-action principle because there is no conservation of mass in this case, the population itself is random. It is unclear that the random process can be driven by Brownian but here we assume a small noise effect for simplicity. Users move over several geographical areas and some of them may carry portable wifi access points. Thus, the network is mobile and random. Each access point may interact with other hotspots in a certain neighborhood of communication and then the information/host propagates over multiple hops. In this setting an increasing rate is observed within the last decade. To capture this phenomenon, we propose an L^α -norm drift model for the rate[46]. The control parameter u_1 of the attacker may represent for example the rate at which the virus attempts transmission over the access points.

4.1 State dynamics models

With explosive growth of mobile devices and the internet of things (IoT), there is an increasing number of vulnerable devices and machines connected to these networks and hotspots so that the mass is not conserved. The population has a tendency the growth (in expectation). To illustrate this we consider an information propagation model where the state is given by

$$dx = \gamma(x)dt + \sigma dB,$$

where $\gamma(x) = \kappa x(1 - \frac{x}{K})$, $\kappa, K > 0$, $x(0) \in (0, K)$. Figure 1 represents the evolution of the state under different noises. In Figure 1 the parameters are $\kappa = 10$, $K = 2$, and $\sigma \in \{0, 1\}$ and different noise terms are plotted. Starting from a initial state value $x(0) = 0.3$, we observe that the population state has tendency to move around 2 within the time interval $[0, 1]$.

4.2 The state needs to be controlled

Due to the presence of malicious attack in the network, there are lot of security, privacy concerns so that the state needs to be controlled [3]. To illustrate the model we introduce an attacker and a defender. Each of them has a control parameter, and has to make a certain decision on those parameters.

$$dx = [\gamma(x) + u_1 - u_2]dt + \sigma dB,$$

u_1 is the attacker control strategy and u_2 is the defender control strategy. In Figure 2 we represent the case where a significant control effort $e = u_2 - u_1 = 0.3$ is injected into the system. We observe that the control affects significantly the state dynamics and help towards a certain goal. For a significant effort

Figure 1: State dynamics with three different noises.

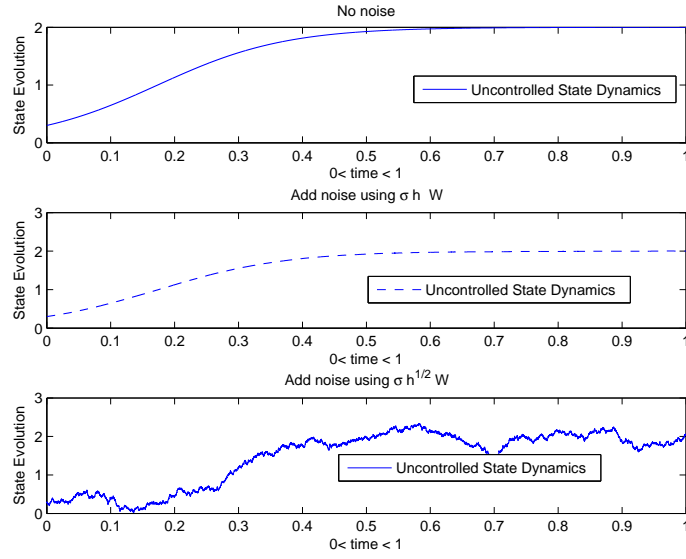
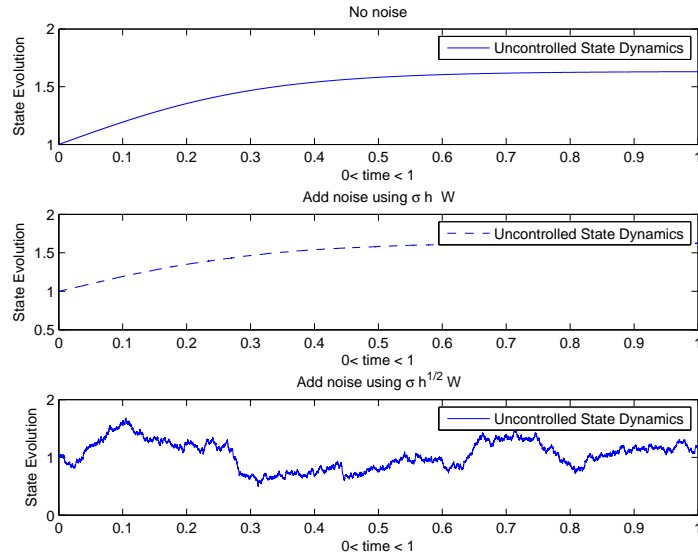


Figure 2: Effect of control effort $e = u_2 - u_1 = 0.3$ in the state.

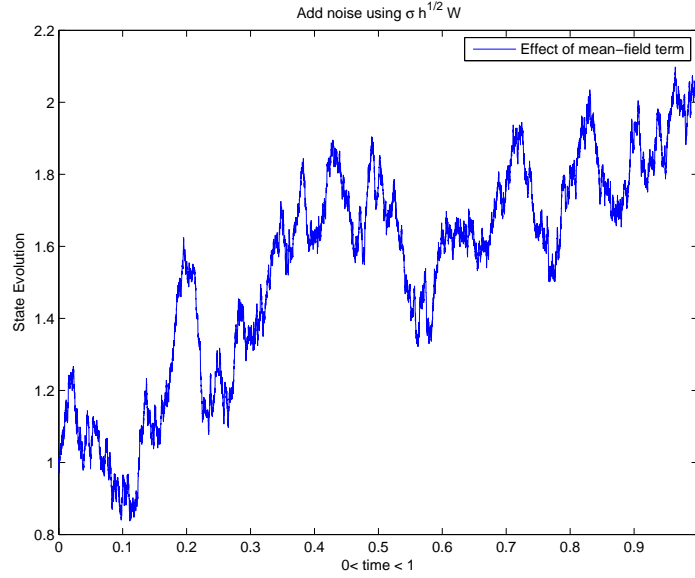


$e = u_2 - u_1 = 0.3$ invested into security, the state of infection can stay below the level 2 starting from level 1. This means that control helps to reduce the infection rate and improve security. However, the location of mobile devices and WiFi hotspots may be important in some cases, specially when local interaction and communications arise. In order to capture this phenomenon we introduce a non-linear behavior via the geographical location distribution and the intensity of interaction at time t as $m(t, dy)$ and we introduce $b(t, x, y, u_1, u_2) = y[\gamma(x) + u_1 - u_2]$, and the state dynamics becomes

$$dx = |\gamma(x) + u_1 - u_2| \left[\int_y y^\alpha m(t, dy) \right]^{\frac{1}{\alpha}} dt + \sigma dB.$$

The variable y can be seen as the intensity of interaction of infected devices/hosts.

Figure 3: Effect of the mean-field term when the effort is $e = u_2 - u_1 = 0.3$. Open-Loop case.



As we can see in Figure 3, the mean-field term $[\int_y y^\alpha m(t, dy)]^{\frac{1}{\alpha}}$ affects significantly the state dynamics in a multiplicative manner. The infection time increases rapidly with the mean-field term. Figure 4 uses a state feedback strategy in the form of $[\int_y y^\alpha m(t, dy)]^{\frac{1}{\alpha}} x$, with $\alpha = 1.2$. We observe that the infection state is significantly reduced compared to the open-loop case of Figure 3. This is illustrated in Figure 5 for several initial states. Hence, it is important to strategically control the mean-field term so that the infected machines remains limited and the damage minimized. In order to do such a minimization we introduce below some objective functions.

Figure 4: Effect of the mean-field term under state and mean-field feedback strategies.

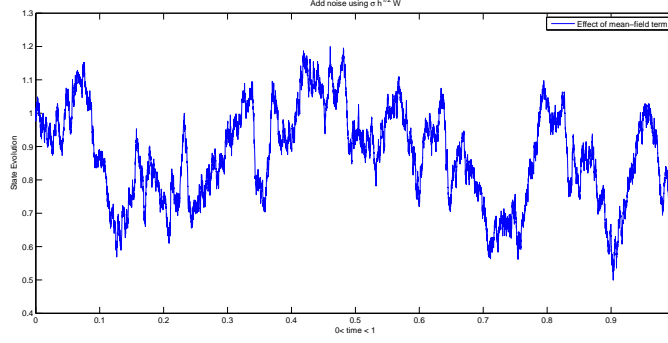
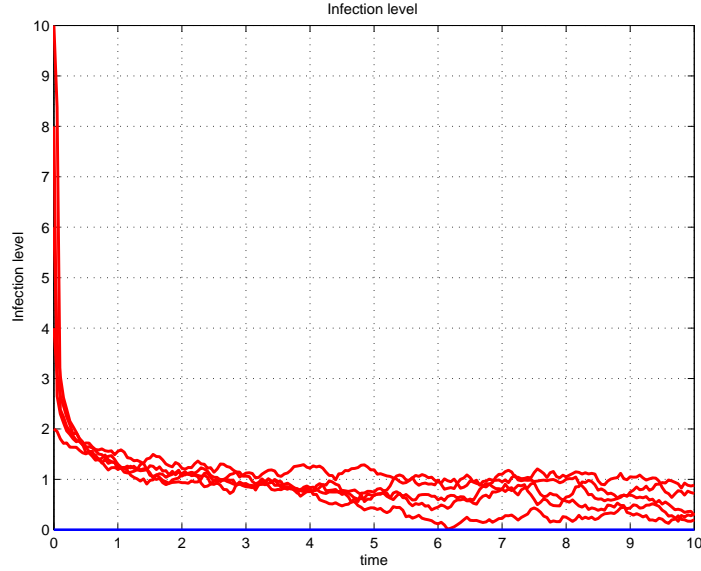


Figure 5: Feedback strategies help to control and maintain the state below a certain range with high probability.

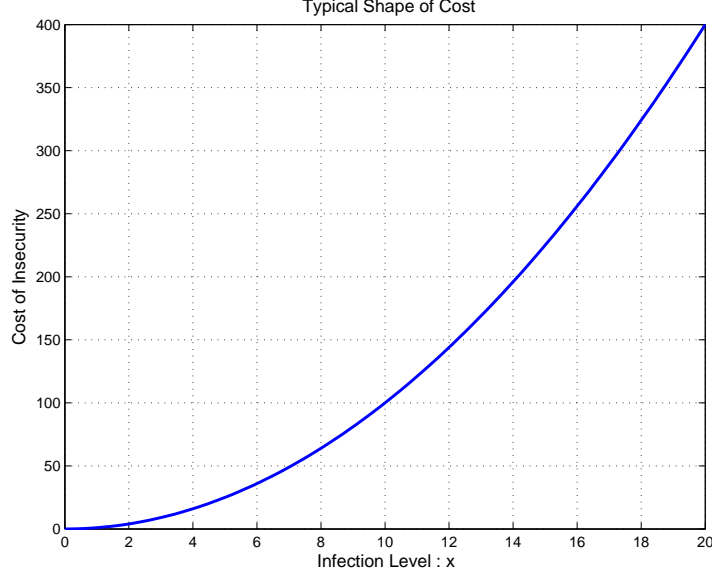


4.3 Objectives

In the context of delay/disruption tolerant networks it can be shown that the delay and the probability of receiving the information have a natural risk-sensitive structure via Poisson arrival rates. However, the attacker and the network defense may not have the same sensitivity when facing the risk. We denote by $\theta_1 = \theta_a$ the attacker risk-sensitivity index and by $\theta_2 = \theta_d$ the defender risk in-

dex. The cost of the attacker is $f_1(x, m, u_1, u_2) = \frac{1}{2}u_1^2$ and its terminal benefit (opposite signed for minimization) is $h_1 = -\frac{c_1}{\alpha}x^\alpha - \frac{\bar{c}_1}{\alpha}m_\alpha$, where m_α denotes the α -th moment of the state process. The goal for the attacker is then to find a tradeoff between the attack effort cost $-\frac{1}{2}u_1^2$ and the benefit $\frac{c_1}{\alpha}x^\alpha + \frac{\bar{c}_1}{\alpha}m_\alpha$.

Figure 6: A typical one step cost function.



The cost of the defender (could be the system administrator) is decomposed into damage cost and security investment loss is $f_2 = \frac{1}{2}x^2 + \frac{1}{2}u_2^2$, $h_2 = \frac{c_1}{\alpha}x^\alpha + \frac{\bar{c}_1}{\alpha}m_\alpha$, $c_1 \geq 0, \bar{c}_1 \geq 0$. For simplicity, the drift is chosen as $\bar{b} = \|b\|_\alpha$ where $b(t, x, y, u_1, u_2) = y[\gamma(x) + u_1 - u_2]$, $\gamma(x) \geq 1$. The control variables are limited to the interval $[0, 1]$ at any time and the diffusion coefficient is system-size dependent: $\sigma_{n(t)} := \frac{\sigma}{n(t)}$ where $n(t)$ is a random variable representing the (active) system size at t . Since there are multiple defense strategies, here we do pull them together in a cooperative manner as an ideal target. However, as observed in practice, the defender may not coordinate their defense strategies due to non-alignment of objectives and/or professional privacy issues. Figure 6 represents a typical instantaneous cost. In Figure 7 we plotted the evolution of a typical cost (random) invested into security over time.

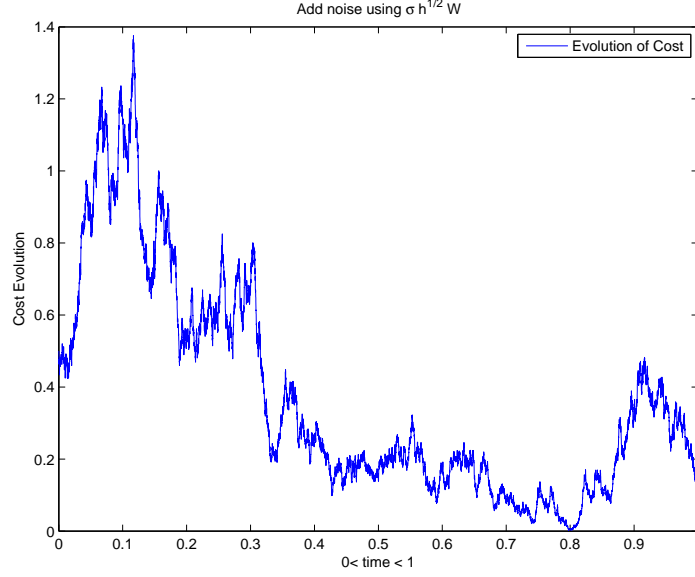
These functions are not bounded, and we cannot use directly the existence results established above. However, we provide the optimality equation for the interior case and derive a risk-sensitive SMP.

$$H_i = m_\alpha^{1/\alpha}[\gamma(x) + u_1 - u_2]p_i + f_i.$$

The attacker's optimal strategy is

$$u_1 = [-m_\alpha^{1/\alpha}p_1]_0^1,$$

Figure 7: Evolution of the cost invested into security.



where $[a]_0^1 := \min(1, \max(0, a))$. The defender's optimal strategy is

$$u_2 = [m_\alpha^{1/\alpha} p_2]_0^1,$$

where p_1, p_2 , solve the risk-sensitive SMP system :

$$dp_i = -\{H_{i,x} + \frac{1}{v_i^\theta} E[v_i^\theta \partial_x H_{i,m}]\} dt + q_i(-\theta_i l_i dt + dB).$$

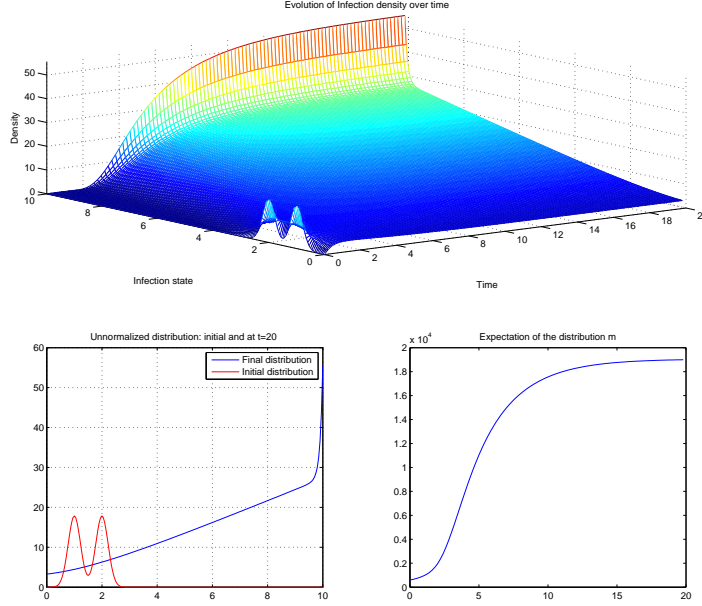
where $H_{i,x} = m_\alpha^{1/\alpha} \gamma'(x) p_i + f_{i,x}$, $f_{1,x} = 0, f_{2,x} = x$, $H_{i,m}(\cdot, t, \xi, m)(x) = \frac{x^\alpha}{\alpha m_\alpha^{\alpha-1}} [\gamma(x) + u_1 - u_2] p_i$, and

$$\begin{aligned} \partial_x H_{i,m}(\cdot, t, \tilde{X}, m)(x) = \\ \frac{x^{\alpha-1}}{m_\alpha^{\alpha-1}} [\gamma(x) + u_1 - u_2] p_i + \frac{x^\alpha}{\alpha m_\alpha^{\alpha-1}} \gamma'(x) p_i. \end{aligned}$$

4.4 Backward-Forward System

$$\begin{cases} dp_1 = -\{H_{1,x} + \frac{1}{v_1^\theta} E[v_1^\theta \partial_x H_{1,m}]\} dt + q_1(-\theta_1 l_1 dt + dB), \\ dp_2 = -\{H_{2,x} + \frac{1}{v_2^\theta} E[v_2^\theta \partial_x H_{2,m}]\} dt + q_2(-\theta_2 l_2 dt + dB), \\ dx^u(t) = \bar{b}(\cdot, t, x^u(t), m^u(t), u(t)) dt + \sigma dB(t), \\ x^u(0) = x_0, \quad m^u(t) := \mathcal{L}(x^u(t)), \end{cases} \quad (35)$$

Figure 8: Mean-field over time, the initial and final distribution, and the evolution of the expected values.



where v_i^θ, l_i and the terminal conditions solve (19) with the optimal strategies

$$(u_1, u_2) = ([-m_\alpha^{1/\alpha} p_1]_0^1, [m_\alpha^{1/\alpha} p_2]_0^1).$$

We investigate (35) numerically under stochastic Euler scheme (also called Euler-Maruyama scheme). We choose $\alpha = 1.2$. We set $\theta_1 = 0.1, \theta_2 = 0.3, c_1 = 0.8 = \bar{c}_1$. The initial distribution is concentrated around two points: 1 and 2. This can be extended to capture a geographical area where the attacks are concentrated in two main countries. In Figure 8, we plot the state distribution over time, the initial and final distribution, and the evolution of the expected values. We observe that the distribution moves progressively towards higher states.

5 Concluding remarks

In this paper we have studied mean-field-type games with a drift that has L^α -norm structure. Although this norm is not differentiable, it is possible to get existence of solutions. We have established relationship between the risk-sensitive SMP and the dual functions. This allow us to verify the risk-sensitive SMP equations. When, the drift is mean-field free, we retrieve the classical

risk-sensitive equations. The work can be extended in several ways. First, the interaction model in the drift $b(t, x, y, u)$ (which is pairwise interaction) can be modified to include k -wise interaction in the form $b(t, x, y_1, \dots, y_k, u)$ with the measure $\prod_{i=1}^k m(t, dy_i)$. This is, in particular, useful for the control of virus spread over network where the interaction involves multiple nodes at a time. Second, the explicit solutions or qualitative analysis of SMP need to be conducted. Third, when $\alpha < 1$, we do not have a norm and the triangular inequality does not hold. In that case, quasi-norm type of inequalities need to be established. We leave these open issues for future research.

References

- [1] Jovanovic, B. (1982): Selection and the Evolution of Industry, *Econometrica* 50, 649-670.
- [2] B. Jovanovic and R. W. Rosenthal (1988). Anonymous sequential games, *Journal of Mathematical Economics*, volume 17, pages 77-87.
- [3] Dario Bauso, Ben Mansour Dia, Boualem Djehiche, Hamidou Tembine , Raul Tempone (2014), Mean-Field Games for Marriage, *PLoS ONE* 9(5): e94933. doi:10.1371/journal.pone.0094933
- [4] A. Bensoussan, , S. C. P. Yam, B. Djehiche and H. Tembine (2014), Risk-Sensitive Mean-Field-Type Control, Preprint.
- [5] Andersson, D. and Djehiche, B. (2010), A maximum principle for SDE's of mean-field type. *Appl. Math. Optim.* 63(3), 341-356.
- [6] A. Bensoussan, K. C. J. Sung, S. C. P. Yam, and S. P. Yung (2012). Linear-quadratic mean-field games. Preprint.
- [7] Bahlali K., Djehiche B. and Mezerdi B. (2007), On the stochastic maximum principle in optimal control of degenerate diffusions with Lipschitz coefficients. *Appl. Math. and Optim.*, 56(3), pp. 364-378.
- [8] Buckdahn, R., B. Djehiche, B. and Li, J. (2011), A general stochastic maximum principle for sdes of mean-field type. *Applied Math. and Optimization*, 64(2), 197-216.
- [9] Buckdahn R. and Li J. and Peng S. (2009), Mean-field backward stochastic differential equations and related partial differential equations, *Stoch. Process. Appl.* ,119(10), 3133-3154.
- [10] Chighoub, F., Mezerdi, B. (2013). A stochastic maximum principle in mean-field optimal control problems for jump diffusions, *Arab Journal of Mathematical Sciences*, Volume 19, Issue 2, July, Pages 223-241.

- [11] El-Karoui, N. and Hamadène, S. (2003), BSDEs and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations. *Stoch. Process. Appl.* (107), 145-169.
- [12] Hafayed, M. (2013), A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes, *International Journal of Dynamics and Control*, December 2013, Volume 1, Issue 4, pp 300-315.
- [13] Hosking, J. (2012), A stochastic maximum principle for a stochastic differential game of a mean-field type. *Appl. Math. and Optim.* 66, pp. 415-454, 2012.
- [14] Jacobson, D.H. (1973), Optimal stochastic linear systems with exponential criteria and their relation to differential games. *Trans. Automat. Control AC*, 18, 124-131.
- [15] Jourdain, B., Méléard, S. and Woyczynski, W. (2008): Nonlinear SDEs driven by Lévy processes and related PDEs. *Alea* 4, 1-29.
- [16] Li, J. (2012), Stochastic maximum principle in the mean-field controls. *Automatica*, 48, pp. 366-373.
- [17] Lim A E B, Zhou X (2005). A new risk-sensitive maximum principle. *IEEE Trans Autom Cont*, 50(7): 958-966.
- [18] Shen, Y. and Siu, T. K. (2013), The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem, *Non-linear Analysis: Theory, Methods and Applications*, Volume 86, July, Pages 58-73.
- [19] Shi, J. and Wu, Z. (2011), A risk-sensitive stochastic maximum principle for optimal control of jump diffusions and its applications. *Acta Mathematica Scientia*, 31(2), pp. 419-433, 2011.
- [20] Shi, J. and Wu, Z. (2012), Maximum Principle for Risk-Sensitive Stochastic Optimal Control Problem and Applications to Finance, *Stochastic Analysis and Applications*, Volume, 30, Issue 6.
- [21] Tembine H. and Zhu Q. and Basar T. (2014). Risk-sensitive mean-field games, *IEEE Transactions on Automatic Control*, 59, 4, pp. 835 - 850, DOI 10.1109/TAC.2013.2289711
- [22] Djehiche B., Tembine H (2015) : A Stochastic Maximum Principle for Risk-Sensitive Mean-Field-Type Control, *IEEE Transactions on Automatic Control*, to appear.
- [23] K. Kometani, H. Shimizu (1975). A study of self-organizing processes of nonlinear stochastic variables, *Journal of Statistical Physics*, December 1975, Volume 13, Issue 6, pp 473-490.

- [24] Djehiche B., Basar T. and Tembine H : Mean-Field-Type Game Theory, Springer, under preparation 2014.
- [25] Haynatzka V.R., Gani J., Rachev S.T (2000): The Spread of AIDS among interactive transmission groups. Mathematical and computer modeling, 32, pp 169-180.
- [26] Tembine, H. (2012). Distributed strategic learning for wireless engineers, *CRC Press, Taylor & Francis*, 496 pages.
- [27] A. Bensoussan (2006), Explicit solutions of linear quadratic differential games, Chapter 2 in Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems, International Series in Operations Research and Management Science Volume 94, 2006, pp 19-34.
- [28] Bensoussan A., Frehse J., Yam S.C.P. (2013): Mean Field Games and Mean Field Type Control Theory, SpringerBriefs in Mathematics, Springer.
- [29] D. Aldous (1985). Exchangeability and related topics. In Hennequin, P., Editor, Ecole d' Ete de Probabilites de Saint-Flour XIII - 1983, Springer-Verlag, Heidelberg. Lecture Notes in Mathematics 1117, pages 1-198.
- [30] J. R. Blum, H. Chernoff, M. Rosenblatt, and H. Teicher (1958). Central limit theorems for interchangeable processes. *Canad. J. Math.*, page 10.
- [31] B. de Finetti (1931). Funzione caratteristica di un fenomeno aleatorio. atti della r. academia nazionale dei lincei, serie 6. memorie, classe di scienze fisiche, matematiche e naturale. 4:251-299.
- [32] E. Hewitt and L. J. Savage (1955). Symmetric measures on cartesian products. *Transactions of the American Mathematical Society*, 80:470- 501.
- [33] P. Kotolenez and Kurtz T (2010). Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type. *Probab. Theory Relat. Fields*, (146):189-222.
- [34] J. Trashorras (2002). Large deviations for a triangular array of exchangeable random variables. *Annales de l'Institut Henri Poincare, Probability and Statistics*, 38: 649-680.
- [35] C. Villani (2009). Optimal transport, old and new. 338 of Grundlehren der mathematischen Wissenschaften, Springer, ISBN 978-3-540-71049- 3.
- [36] Cisse A.K, Tembine H (2014): Cooperative Mean-Field Type Games, Proc. of 19th World Congress The International Federation of Automatic Control Cape Town, South Africa. August 24-29.
- [37] Lauriere Mathieu, Pironneau Olivier (2014), Dynamic programming for mean-field type control, *C. R. Acad. Sci. Paris, t. , Serie I*.

- [38] Realtime web attack monitoring (2014) at <http://www.akamai.com/html/technology/real-time-web-monitor.html>
- [39] James Bergin, Dan Bernhardt (1992), Anonymous Sequential Games with Aggregate Uncertainty, *Journal of Mathematical Economics* Volume 21, Issue 6, Pages 543-562
- [40] Jean-David Benamou and Yann Brenier (2000). A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375-393.
- [41] J. D. Benamou and Y. Brenier (2001). Mixed L2-Wasserstein optimal mapping between prescribed density functions. *J. Optim. Theory Appl.*, 111(2):255-271.
- [42] H. Tembine (2014), Risk-Sensitive Mean-Field-Type Games with L^p -norm Drifts, *Automatica*, to appear.
- [43] Jean-Michel Lasry and Pierre-Louis Lions (2007). Mean field games. *Jpn. J. Math.*, 2(1):229-260.
- [44] M.Y. Huang, R.P. Malhame and P.E. Caines (2006). Large Population Stochastic Dynamic Games: Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle, Special issue in honor of the 65th birthday of Tyrone Duncan, *Communications in Information and Systems*. Vol 6, Number 3, pp 221-252.
- [45] Djehiche B. and Tembine H. (2014): A Stochastic Maximum Principle for Risk-Sensitive Mean-Field-Type Control under Partial observation. *Stochastic of Environmental and Financial Economics*, CAS SEFE Proceedings.
- [46] H. Tembine (2014): Energy-Constrained Mean-Field Games in Wireless Networks, *Strategic Behavior and the Environment*: Vol. 4: No. 2, pp 187-211. <http://dx.doi.org/10.1561/102.000000040>

Appendix

A Mean-field convergence

The extension of (i) the law of large numbers, (ii) central limit theorem and (iii) large deviation principle, from independent random variables to sequences of indistinguishable random variables has drawn the attention of a number of researchers ever since the appearance in Blum, Chernoff and co-authors [30]. Below we present some well-known results and explain how they can be used in the McKean-Vlasov context with the L^α -norm.

A.1 Indistinguishability

The notion of indistinguishability (or exchangeability or interchangeability) is introduced in order to discuss the existence of a limiting measure and mean-field convergence of the empirical measure of virtual particle states in the framework of de Finetti-Hewitt-Savage [31, 32, 29, 33, 34].

Let \mathcal{X} be a separable complete and metrizable topological space (Polish space).

Definition 3 (Indistinguishability). *A collection $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ of \mathcal{X} -valued random variables/processes, is indistinguishable (or exchangeable) if the joint law is invariant by permutation over the index set $\{1, \dots, n\}$, i.e., for any permutation σ over the set $\{1, 2, \dots, n\}$, one has*

$$\mathcal{L}(x_{(1)}, x_{(2)}, \dots, x_{(n)}) = \mathcal{L}(x_{\sigma(1)}, \dots, x_{\sigma(n)}), \quad (36)$$

where $\mathcal{L}(X)$ denotes the law of the random variable X . An infinite family of random variables/processes $(x^{(1)}, x^{(2)}, \dots)$ is indistinguishable if every finite n , the family $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ is indistinguishable.

This says that the order (position) of the random variable in the family does not change the joint distribution. From (36) we also have that, for any measurable operator O ,

$$\begin{aligned} & \mathcal{L}\left(x_{(1)}, x_{(2)}, \dots, x_{(n)}, O\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_{(i)}}\right)\right) \\ &= \mathcal{L}\left(x_{\sigma(1)}, \dots, x_{\sigma(n)}, O\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_{(i)}}\right)\right), \end{aligned} \quad (37)$$

where we do not permute the last component.

For indistinguishable random variables/ processes, the convergence of the empirical measure $m_n := \frac{1}{n} \sum_{i=1}^n \delta_{x_{(i)}}$ has been widely studied. This sits at the intersection between group theory and probability theory. The symmetry group properties have been used to derive some properties of the distributions of the processes. The *de Finetti-Hewitt-Savage* theory provides the mean-field convergence of such a measure-valued process. When studying convergence of measures, an important issue is the choice of probability metric. In order to measure the gap between two probability measures, we introduce the Wasserstein (Vasershtein) metric (also called Monge-Kantorovich metric) d_α of order $\alpha \geq 1$.

Definition 4 (Wasserstein).

$$\begin{aligned} d_\alpha^\alpha(\mu, \nu) &= \inf \left\{ \int_{(x,y)} d_0(x, y)^\alpha \gamma(dx, dy); \right. \\ & \left. \gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{X}), \gamma_x = \mu, \gamma_y = \nu \right\}, \end{aligned}$$

where γ_x denotes the marginal with the respect to the x -component, where d_0 is a reference metric on \mathcal{X} (such a metric exists because \mathcal{X} is assumed to be metrizable).

The famous Kantorovich-Rubinstein 1958 theorem gives a dual representation of d_1 in terms of a Lipschitz-Bounded metric:

$$d(\mu, \nu) := d_1(\mu, \nu) = \sup \left\{ \int \phi d(\mu - \nu); \|\phi\|_{Lip} \leq 1 \right\},$$

where $\|\phi\|_{Lip} = \|\phi\|_\infty + \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{d_0(x, y)}$ is the Lipschitz-norm of ϕ . It can be shown that d_α is a metric (a "true" distance in a topological sense), i.e., it satisfies the axioms of a metric. For Polish spaces \mathcal{X} , the Wasserstein distance d_1 is known to metrize the weak topology over \mathcal{X} . As stated in Villani's book [35] the Wasserstein distance has the following properties: for any $1 \leq \alpha < +\infty$,

$$\lim_n d_\alpha(m_n, m) = 0$$

implies, in particular, that

- m_n converges to m in distribution (weak convergence of probability measures) i.e.,

$$E_{m_n}[\phi] := \int \phi dm_n \rightarrow E_m[\phi],$$

as $n \rightarrow +\infty$, for any measurable bounded and Lipschitz functions ϕ .

- $\int d_0^\alpha(x, y) m_n(dy) < +\infty$ for some $x \in \mathcal{X}$.

Thanks to these nice properties, the Wasserstein distance d_α is an appropriate candidate for the convergence of the empirical measure in the weak sense.

Theorem 1 (de Finetti-Hewitt-Savage). *Let $x_{(1)}, x_{(2)}, \dots$, be an indistinguishable sequence of \mathcal{X} -valued random variables, where \mathcal{X} is a Polish space. Then, there is a $\mathcal{P}(\mathcal{X})$ -valued random measure m such that*

$$m = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{x_{(i)}}, \text{ almost surely,}$$

where $\mathcal{P}(\mathcal{X})$ denotes the space of probability measures on \mathcal{X} . Conditioned on m , the random variables $x_{(1)}, x_{(2)}, \dots$ are i.i.d with distribution m , that is, for each measurable bounded function ϕ ,

$$\begin{aligned} & \mathbb{E}(\phi(x_{(1)}, x_{(2)}, \dots, x_{(k)}) \mid m) \\ &= \int \phi(y^1, \dots, y^k) m(dy^1) \dots m(dy^k). \end{aligned}$$

In addition, if the moments of $x_{(i)}$ are finite then

$$d_1 \left(m, \frac{1}{n} \sum_{i=1}^n \delta_{x_{(i)}} \right) \leq \frac{C_1}{\sqrt{n}} = O \left(\frac{1}{\sqrt{n}} \right),$$

where $C_1 > 0$ and $d = d_1$ denotes the Wasserstein metric of order one.

Note that the convergence in Theorem 1 is in the weak sense since the Monge-Kantorovich distance d_1 metrizes the weak topology. Theorem 1 has been proved by de Finetti (1931, [31]) for infinite binary sequences and has been extended by Hewitt and Savage (1955, [32]) to continuous and compact state spaces. A simple and elegant proof can be found in Aldous (1985, [29]), pp. 18-22, for the general state space. The rate of convergence for Monge-Kantorovich distance is obtained following the line of the law of large numbers of interacting systems. Theorem 1 was initially used for static (time-independent) maps. Then, several applications in mathematical physics and biology, with dynamical models came into the picture. These are dynamically interacting particles, genes, molecules or nodes. Theorem 1 was then extended to the dynamical case in at least two ways: (i) path wise (up to a certain time step T), (ii) at each time step t .

A.2 Large Deviation Principle for m_n

We say that for any time t , the probability measures $(m_n(t))_{n \geq 0}$ on a topological space obeys a Large Deviation Principle with rate functions $(I(t, \cdot))$ and in the scale $(a_n)_n$ if $(a_n)_n$ is a real-valued sequence satisfying $a_n \rightarrow \infty$ and I is a non-negative, lower semicontinuous function such that

$$\begin{aligned} - \inf_{x \in \text{int}(B)} I(t, x) &\leq \liminf_n \frac{1}{a_n} \log m_n(t, B) \\ &\leq \limsup_n \frac{1}{a_n} \log m_n(t, B) \leq - \inf_{x \in \text{cl}(B)} I(t, x), \end{aligned}$$

for any measurable set B , whose interior is denoted by $\text{int}(B)$ and closure by $\text{cl}(B)$. If the level sets $\{x : I(t, x) \leq \beta\}$ are compact for every $\beta < +\infty$, $I(t, \cdot)$ is called a *good rate function*. We introduce $\tilde{H}(\cdot | \cdot)$ as the relative entropy function (defined also above)

$$\tilde{H}(\mu | \nu) := \int \log\left(\frac{d\mu}{d\nu}\right) d\mu,$$

if μ is absolutely continuous with the respect to ν and $+\infty$ otherwise.

The main advantage of having this type of result is the decay of $m_n(t, B)$ as n gets large. Basically, when the two limits are identical, $m_n(t, B)$ is the order of $e^{-a_n R}$ where $R = \inf_{x \in B} I(t, x) > 0$. As a consequence, the weak convergence from Theorem 1 and central limit theorems can be derived from these inequalities.

The next result provides a large deviation principle result [35].

Theorem 2. *Assume that initially $m_n(0)$ follows a large deviation principle with rate $I(0, m(0))$ on the set of probability measures $\mathcal{P}(\mathcal{X})$. Then $(m_n(t, \cdot))_{t \in [0, T]}$ follows a large deviation principle on the set of cadlag functions from $[0, T]$ to $\mathcal{P}(\mathcal{X})$ with good rate*

$$\begin{aligned} I(t, m(t)) &= \\ \begin{cases} I(0, m(0)) + \int_0^1 \tilde{H}(\dot{m}(s) | m(0)) \text{ if } \dot{m}(s) \text{ a.c.} \\ +\infty \text{ otherwise} \end{cases} \end{aligned}$$

where a.c. means absolutely continuous.

Definition 5. Consider two processes $(x(t))_{t \in [0, T]}$ and $(y(t))_{t \in [0, T]}$ and set

$$D_{T, \alpha}(\mu, \nu) := \inf \left\{ \left(E \left[\sup_{t \in [0, T]} d_0(x(t), y(t))^\alpha \right] \right)^{1/\alpha} ; \right. \\ \left. \mathcal{L}(x) = \mu, \mathcal{L}(y) = \nu \right\}.$$

We now prove Theorem 1 in several steps:

A.3 Existence of solution to the state equation

We start with the existence of solution. To prove existence of a solution with the respect to the Wasserstein distance, we adopt a contraction-type of approach. Then, we construct a Cauchy sequence (L^α space which is a Polish space). Consequently, the solution is almost unique. Consider the SDE given by

$$x(t) = x(0) + \int_0^t \sigma(\cdot, s, x(s)) dB(s) + \\ \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x(s), y, u) m(s, dy) \right]^{1/\alpha} ds =: RH(x)[m].$$

Then one gets a fixed-point stochastic equation in m : $m = \mathcal{L}(x(t)) = \mathcal{L}(RH(x)[m])$.

Consider two measures $m_1(s, dy)$ and $m_2(s, dy)$ such that $D_{T, \alpha}(m_i, m_0) < +\infty$ then

$$D_{t, \alpha}(\mathcal{L}(RH(x)[m_1]), \mathcal{L}(RH(x)[m_2])) \\ \leq 2Le^{2Lt} \int_0^t D_{s, \alpha}(m_1, m_2) ds,$$

for any $t \in [0, T]$. Let $c > 0$, $\Psi(f)(t) \leq c \int_0^t f(s) ds$, $t \in [0, T]$ for some operator Ψ . and $M = \sup_{s \in [0, T]} f(s)$.

$$\Psi^2(f)(t) \leq c \int_0^t \Phi(f)(s) ds = c \int_0^t [cMs] ds = c^2 \frac{t^2}{2} M.$$

Then, by induction, $\Phi^k(f)(t) \leq \left(c^k \frac{T^k}{k!} \right) M$, and

$$D_{t, \alpha}(\mathcal{L}^{k+1}(\mu_0), \mathcal{L}^k(\mu_0)) \leq C_T^k \frac{T^k}{k!} D_{T, \alpha}(\mathcal{L}(\mu_0), \mu_0) < +\infty,$$

where $C_T = 2Le^{2LT}$, and where we have used the Lipschitz continuity L of b and the Minkowski inequality for $p \geq 1$. By summing up over natural numbers k , one gets $\sum_{k \geq k_0} D_{t, \alpha}(\mathcal{L}^{k+1}(\mu_0), \mathcal{L}^k(\mu_0)) < +\infty$. Thus, $(\mathcal{L}^k(\mu_0))_{k \geq k_0}$, is a Cauchy sequence with the respect to the metric $D_{T, \alpha}$, for $k_0 \geq 1$. Since we are in a complete metric space, this sequence converges to some fixed-point (say, m). Then $m = \mathcal{L}(RH(x)[m]) = \mathcal{L}(m)$ is the unique fixed-point, solution of the SDE.

A.4 Pathwise mean-field convergence

Consider n independent random process $x_{i,n}^*$ satisfying

$$\begin{aligned} x_{i,n}^*(t) &= x_{i,n}^*(0) + \int_0^t \sigma(\cdot, s, x_{i,n}^*(s)) dB(s) \\ &+ \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m^*(s, dy) \right]^{1/\alpha} ds, \end{aligned}$$

$m^*(t, \cdot) = \mathcal{L}(x^*(t))$. and the particle representation

$$\begin{aligned} x_{i,n}(t) &= x_{i,n}(0) + \int_0^t \sigma(\cdot, s, x_{i,n}(s)) dB(s) \\ &+ \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}(s), y, u) m_n(s, dy) \right]^{1/\alpha} ds, \end{aligned}$$

where $m_n^*(t, \cdot) = \frac{1}{n} \sum_{i=1}^n \delta_{x_{i,n}^*(t)}$.

It suffices to prove the statement for the coefficient with mean-field term. We take the difference between the two drift terms.

$$D = \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m^*(s, dy) \right]^{1/\alpha} ds \quad (38)$$

$$- \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}(s), y, u) m_n(s, dy) \right]^{1/\alpha} ds \quad (39)$$

We decompose D into three separate terms as follows:

$$\begin{aligned} D &= \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m^*(s, dy) \right]^{1/\alpha} ds \\ &- \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m_n^*(s, dy) \right]^{1/\alpha} ds \quad (40) \end{aligned}$$

$$\begin{aligned} &+ \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m_n^*(s, dy) \right]^{1/\alpha} ds \\ &- \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m_n(s, dy) \right]^{1/\alpha} ds \quad (41) \end{aligned}$$

$$\begin{aligned} &+ \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m_n(s, dy) \right]^{1/\alpha} ds \\ &- \int_0^t \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}(s), y, u) m_n(s, dy) \right]^{1/\alpha} ds \quad (42) \end{aligned}$$

$$D = I_1 + I_2 + I_3 \quad (43)$$

The first term I_1 (see (40)) deals only with i.i.d random variables. Therefore, the convergence for that part is classical. For the second term I_2 , we use the triangular inequality for $\alpha \geq 1$. By Lipschitz continuity of b , we get $|I_2| \leq \frac{L}{n^{\frac{1}{\alpha}}} \int_0^t [\sum_{i=1}^n \|x_{i,n}^*(s) - x_{i,n}(s)\|^\alpha]^{1/\alpha} ds$. By Lipschitz continuity of b , $|I_3| \leq \int_0^t |x_{i,n}^*(s') - x_{i,n}(s)| ds$.

By Hölder inequality

$$\begin{aligned}
& \left\{ E \left[\sup_{s \in [0, T]} |x_{i,n}^*(s) - x_{i,n}(s)|^\alpha \right] \right\}^{1/\alpha} \\
& \leq \int_0^T \{E \|Y_{n,s}\|^\alpha\}^{1/\alpha} ds \\
& + \frac{L}{n^{\frac{1}{\alpha}}} \int_0^T \left[E \sup \left\| \sum_{j=1}^n |x_{j,n}^*(s) - x_{j,n}(s)| \right\|^\alpha \right]^{1/\alpha} ds \\
& + L \int_0^T \{E \left[\sup_{s \in [0, t]} |x_{i,n}^*(s) - x_{i,n}(s)|^\alpha \right]\}^{1/\alpha} dt
\end{aligned} \tag{44}$$

where

$$\begin{aligned}
Y_{n,s} &= \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m^*(s, dy) \right]^{1/\alpha} \\
&\quad - \left[\int_y |b|^\alpha(\cdot, s, x_{i,n}^*(s), y, u) m_n^*(s, dy) \right]^{1/\alpha}
\end{aligned} \tag{45}$$

Summing over i and using Gronwall Lemma yields

$$\begin{aligned}
& \{E \left[\sup_{s \in [0, T]} |x_{i,n}^*(s) - x_{i,n}(s)|^\alpha \right]\}^{1/\alpha} \\
& \leq 2Le^{2LT} \int_0^T \{E \|Y_{n,s}\|^\alpha\}^{1/\alpha} ds.
\end{aligned}$$

which provides a convergence rate of \sqrt{n} since

$$\sup_n \sqrt{n} \left\{ E \sup_{s \in [0, T]} \|Y_{n,s}\|^\alpha \right\}^{1/\alpha} < +\infty.$$

This result shows that when $\alpha \geq 1$, and when the initial distributions of the virtual particles are mutually independent, with same distribution as $x(0)$, then the particle interaction model with fixed control u soon destroys that independence through the empirical measure m_n . But, for a given finite time t , when the number of particle becomes large, the mean-field convergence implies that the distributions become approximately independent again conditioning on $m(t, \cdot)$, so that independence is still retained. This is called *propagation-of-chaos*. Note

that this result is limited to finite horizon. For long-term behavior one needs to study the asymptotics (infinite horizon in time) of the SDEs in order to derive propagation (or non-propagation) of chaos property.

Hamidou Tembine received his M.S. degree in Applied Mathematics from Ecole Polytechnique in 2006 and his Ph.D. degree in Computer Science from University of Avignon in 2009. His current research interests include evolutionary games, mean field stochastic games and applications. In 2014, Tembine received the IEEE ComSoc Outstanding Young Researcher Award for his promising research activities for the benefit of the society. He was the recipient of 5 best paper awards in the applications of game theory. Tembine is a prolific researcher and holds several scientific publications including magazines, letters, journals and conferences. He is author of the book on distributed strategic learning for engineers (published at CRC Press, Taylor & Francis 2012), and co-author of the book Game Theory and Learning in Wireless Networks (Elsevier Academic Press). Tembine has been co-organizer of several scientific meetings on game theory in networking, wireless communications and smart energy systems. He is a senior member of IEEE.